# Designing Effective Sampling Strategies in Biomedical Studies: is Selection Bias Always Bad?

Fulvia Mecatti - *University of Milano-Bicocca*

The potential of the Sampling Statistics approach for designing biomedical studies is discussed. The focus will be shifted from representativness, as a classical argument of validity in observational study, to the effectiveness of the sampling strategy as tailored upon specific, usually challanging features of the study at hand. This perspective allows broader objectives of the study to be included in the sampling design, beyond inference on population quantities of interest toward making an impact on the population itself. On the other hand such effectiveness may hardly be pursued by means of traditional equal-probability sampling designs granting representative samples. Instead we will consider sampling designs allowing for over-representing into the sample specific segments of the population, as particularly significant for the study, and under-represent other less relevant population units. Indeed this figures a selection bias, although purposively introduced into the sample selection and thus to be adjusted afterword, at the estimation stage via the appropriate weighting system. The sample would be still random and, as a result, the released estimates would be reprodicible, unbiased and with accuracy estimable from the same data. A sampling strategy is proposed which combines a *sequential* units' selection with an *adaptive* data collection. The adaptive component allows tailoring the data collection to peculiar aims and challenges of the study, as for instance the over-sampling of positive cases of a rare desease. At the same time the sequential component would provide a flexible framework for controlling specific constraints of the study, such as clinical, ethical and logistical, at both the design level and in real time sampling.The sampling strategy will be illustrated by means of two observational studies in epidemiology and in oncologycal surgery.

**Keywords**: Adaptive Sampling; List-sequential Sampling; Spatial pattern; Unbiased estimation; Unequal Selection Porbability.

## Essential References

Thompson S.K. (2017) Adaptive and network sampling for inference and interventions in changing populations. *J. of Survey Statistics and Methodology* 5, 1-21

Bondesson L. and Thorburn D. (2008) A list-sequential sampling method suitable for real-time sampling. *Scand. J. of Statistics* 35, 466-483

Floyd S., Sismanidis C., Mecatti,F., Floyd K. and Yamada N. (2011) *Sampling design.* In *Tuberculosis Prevalence Surveys: a handbook* (pp. 51-80). Geneva: World Health Organization

Mecatti F. and Furfaro E. (2018) Sequential Adaptive strategy for population-based sampling of a rare and clustered disease. *Working Paper*

Girigliano C., Nonis A., Briganti A., Bonetti M. and Di Serio C. (2018) Effect of the number of removed lymph nodes on prostate cancer recurrence and survival: evidence from an observational study *???*